

Subhrato Som

Philadelphia, Pennsylvania 19104

📞 215-240-9510 ✉ ss5654@drexel.edu [in linkedin.com/in/subhratosom/](https://www.linkedin.com/in/subhratosom/) github.com/Subhrato20

Education

Drexel University

Master's of Science in Computer Science

Sep 2023 - May 2025

Philadelphia, PA, USA

Vellore Institute of Technology

Bachelor of Technology in Computer Science and Engineering

Jul 2019 - Jun 2023

Chennai, TN, India

Technical Skills

Machine Learning: PyTorch, TensorFlow, Transformers, LangChain, Scikit-learn, LLM Fine-tuning

Cloud & MLOps: AWS SageMaker, EC2, Lambda, Docker, Kubernetes, CI/CD Pipelines

Languages: Python, JavaScript, Java, C++

ML Deployment: REST APIs, Model Optimization, Vector Databases, Milvus

Data Engineering: PostgreSQL, MongoDB, Data Preprocessing, ETL Pipelines

Experience

AI Developer (COOP)

Jun 2024 - Dec 2024

Mathtech, Inc

- Designed and deployed production-grade Retrieval-Augmented Generation (RAG) system using transformer architectures, improving document retrieval accuracy by 32% through fine-tuning embedding models
- Implemented end-to-end ML pipeline with AWS SageMaker for distributed training and model deployment, achieving 99.8% uptime through Kubernetes orchestration
- Optimized vector search performance using Milvus vector database, reducing latency by 40% through query optimization and indexing strategies
- Developed CI/CD pipelines with Jenkins for automated model retraining and deployment, integrating unit testing and model validation checks

Research Assistant

Feb 2024 - Present

Social-NLP Lab, CCI Drexel University

- Pioneered multimodal analysis of social media content using Vision-Language Models (VLMs), achieving 89% accuracy in cross-modal semantic alignment
- Developed novel preprocessing pipelines for unstructured social data, handling 500k+ text-image pairs through distributed processing

Projects

LLM-based Disaster Risk Analytics | Philly Codefest 2024, Chubb Challenge Winner

Apr 2024

- Integrated transformer models with geospatial data analysis, developing ensemble models that improved disaster prediction accuracy by 27% over baseline
- Deployed multi-modal Flask API serving ML models and LLM-generated insights through AWS EC2, handling 150+ concurrent requests with Docker containerization

Medical Image Understanding System | Vellore Institute of Technology (Capstone Project)

Nov 2022 - May 2023

- Developed hybrid CNN-Transformer architecture for joint image classification and report generation, achieving SOTA 89.1% diagnostic accuracy on NIH Chest X-ray dataset
- Implemented attention mechanisms for cross-modal alignment between radiographic findings and textual reports, reducing hallucination by 41%

BERT-based Biomedical QA System | Vellore Institute of Technology

Feb 2022 - Apr 2022

- Fine-tuned BERT models on biomedical corpus using contrastive learning, improving QA accuracy from 76% to 84.1% on domain-specific benchmarks
- Designed knowledge-aware attention layers incorporating UMLS medical ontologies, enhancing model's clinical reasoning capabilities

Activities

Chubb Challenge Winner, Philly Codefest 2024

Apr. 2024

React Summit, GitNataion

June 2022

Open Source Contributor, GirlsScript Foundation

Sep. 2021 - Dec 2021

Student Expert, Postman

Aug. 2021 - Jul 2023

Certifications

Advanced NLP with Transformers, Coursera

Mar 2024

AWS Machine Learning Engineer, Microsoft Learn

Dec 2023

Deep Learning Specialization, LinkedIn Learning

Oct 2022